

Comparación BD NoSQL

Yaneth Virginia Aquino Huallpa, Arlyn Cotrado Coaquira, Sharon Sosa Bedoya, and Marlon Villegas Arando
Universidad Privada de Tacna \Facultad de Ingeniería \Escuela Profesional de Ingeniería de Sistemas

Resumen

En el presente artículo se relata la comparación hecha de bases de datos NoSQL, describiéndolas y analizando su importancia, así como las definiciones de tipos de bases de datos NoSQL, con el fin de proporcionar un punto de partida para los trabajos en esta área. Y también la creación de una base de datos, inserción y consultas de datos NoSQL mediante Docker.

Palabras clave: NoSQL, Bases de datos, Docker.

Abstract

In the present article the comparison made of NoSQL databases is described, describing them and analyzing their importance, as well as the definitions of NoSQL database types, in order to provide a starting point for work in this area. And also the creation of a database, insertion and queries of NoSQL data through Docker.

Keywords: NoSQL, Databases, Docker.

I. INTRODUCCIÓN

Día a día el manejo de la información se hace más complejo; diferentes factores hacen que las personas involucradas en el área busquen tecnologías que le ayuden con este problema. Las bases de datos relacionales son las más comunes, pero en los últimos años ha aumentado el interés por las bases de datos NoSQL (Not only SQL), un nuevo conjunto de tecnologías que pueden contribuir al manejo de información.

Por lo anterior, el presente documento hace una revisión de las tecnologías NoSQL, haciendo posible hacer una comparación.

El resto de este artículo está organizado de la siguiente manera. En la Sección 2 se muestra los materiales y métodos usados para el desarrollo de este artículo. La Sección 3 se explican los resultados. Y finalmente, las conclusiones están en la Sección 4.

II. MATERIALES Y MÉTODOS

A. Materiales

- Virtualización activada en el BIOS
- Docker Desktop
- Windows 10 64bit: Pro, Enterprise o Education, con al menos 4GB de RAM.

B. Métodos

- Se utilizó como material artículos y libros relacionados a la base de datos NoSQL y sus tipos, así como páginas web.

III. MARCO TEÓRICO

A. Base de datos

- Una base de datos es una colección de datos organizados según un determinado criterio
- Estos datos se pueden leer, crear, actualizar y borrar
- También existen motores de base de datos que nos permiten hacer todas estas operaciones de forma más fácil

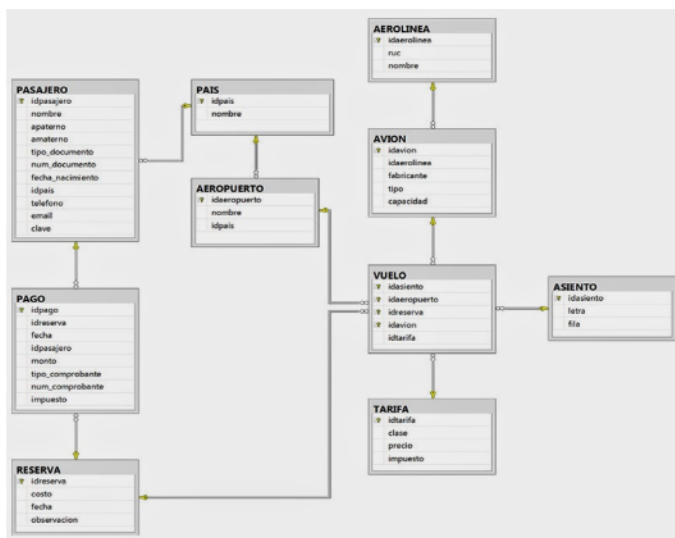
B. Tipos de base de datos

- Existen distintos tipos de bases de datos que se utilizan para solucionar distintos tipos de problemas
- Dentro de la gran familia de bases de datos podemos encontrar las del tipo base de datos relacionales y las no relacionales
- Las bases de datos relacionales se conocen generalmente como las SQL
- Las no relacionales se conocen como NoSQL

- Cada tipo de base de datos tiene beneficios y contras a la hora de almacenar, leer, actualizar o borrar los datos

C. Bases de datos relacional

- Desde su definición vemos que una base de datos puede ser relacional si cumple con algo conocido como el modelo relacional.
- En la definición de modelo relacional nos podemos quedar con la idea de tablas que tienen columnas para describir los datos que están relacionados entre si.
- Ejemplo modelo relacional:



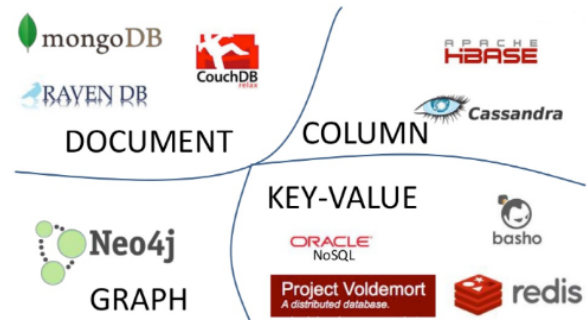
D. Modelo NoSQL

Este tipo de Base de datos es la respuesta a la necesidad de gestionar volúmenes masivos de información, de tal manera que surge la base de datos NoSQL, término que fue acuñado a finales de los años 90 y que engloba todas las tecnologías de almacenamiento estructurado que no cumplen el esquema relacional, mayormente usado.

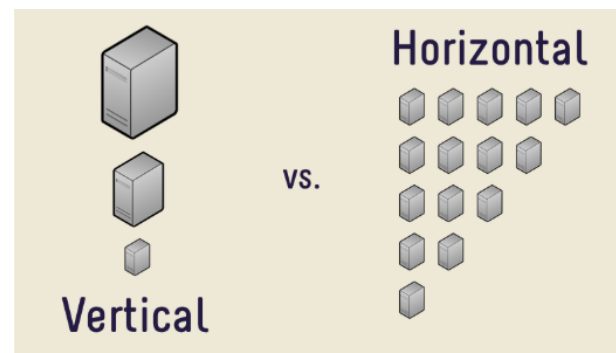
La cantidad de información manejada por las comunidades, redes sociales, buscadores, y muchos otros proyectos en el ámbito de la Web es enorme y abrumadora, lo que ha hecho que surjan nuevas arquitecturas de almacenamiento de información, que deben ser de alto rendimiento, escalables y distribuidas.

- Se conoce como NoSQL (Not Only SQL) al grupo de bases de datos que no son relacionales

- Dentro de esta clasificación se encuentran las bases de clave/valor, orientadas a documentos, grafos, de grandes columnas .



- Este tipo de bases de datos escala de forma horizontal
- Podemos utilizar muchas máquinas chiquitas para crecer y satisfacer las necesidades de los negocios actuales.



IV. RESULTADOS

A. Creacion de base de datos NoSQL con MongoDB

- MongoDB es una base NoSQL orientada a documentos
- Permite guardar documentos en formato de JSON
- Tiene esquema flexible, es decir que podemos cambiar la estructura de nuestros documentos sin ningún problema
- MongoDB está preparado para escalar fácilmente de manera horizontal
- Dado que aprendimos ECMAScript vamos a utilizar un motor de base de datos que nos permite seguir utilizando este lenguaje para guardar nuestros datos.

B. Instalar MongoDB en Docker

- Ingresar sus credenciales creadas en Docker Hub para iniciar sesión en el aplicativo. Ubicar la aplicación PowerShell, ejecutarla como Administrador. En la ventana de comandos de PowerShell escribir lo siguiente.



- Para instalar MongoDB primero tenemos que ejecutar el siguiente código.

```
PS C:\> docker run --name mongobongo -v /home/dac/Downloads/mongodb:/data/db -d mongo:3.5
```

- Instalamos la versión completa de MongoDB

```
PS C:\> docker run --name mongobongo -v /home/dac/Downloads/mongodb:/data/db -d mongo:3.5
Unable to find image 'mongo:3.5' locally
B.5: Pulling from library/mongo
d13d02fa248d: Pull complete
bc8e2652ce92: Pull complete
3cc856886986: Pull complete
c319e9ec4517: Pull complete
b4cbf888f94: Pull complete
e42648c7b452: Pull complete
bfce988bf680: Pull complete
bd3ae40a4e35: Pull complete
dbdbb4a86824: Pull complete
d4624d0639de: Pull complete
Digest: sha256:82e1da5703ed53b03b8f7cac9e531b114d7493476e04b5d624e968f38bb97f07
Status: Downloaded newer image for mongo:3.5
ce711c6ddc679c292e7a793c677dc8aaa3b7b837ca0f5b19e4f689f6cd2e8754
PS C:\>
```

- Verificar que el contenedor se este ejecutando correctamente mediante el comando:

PS C:\> docker ps	CONTAINER ID	IMAGE	COMMAND	CREATED	STATUS	PORTS
ce711c6ddc67	mongo:3.5	"docker-entrypoint.s..."	35 minutes ago	Up 35 minutes	27017/tcp	

- Proceder a verificar la imagen con el siguiente comando:

PS C:\> docker images	REPOSITORY	TAG	IMAGE ID	CREATED	SIZE
docker4u/nsenter-dockerd	latest	2f1c802f322f	8 months ago	187KB	
mongo	3.5	a2ab5e364821	19 months ago	371MB	
store/oracle/database-enterprise	12.2.0.1	12a599cd0528	22 months ago	3.44GB	

- Descargamos un fichero JSON para subir luego a una base MongoDB:

PS C:\> docker images	REPOSITORY	TAG	IMAGE ID	CREATED	SIZE
docker4u/nsenter-dockerd	latest	2f1c802f322f	8 months ago	187KB	
mongo	3.5	a2ab5e364821	19 months ago	371MB	
store/oracle/database-enterprise	12.2.0.1	12a599cd0528	22 months ago	3.44GB	

- MongoDB está preparado para escalar fácilmente de manera horizontal

```
PS C:\> docker run --name mongobongo -v /home/dac/Downloads/mongodb:/data/db -d mongo:3.5
Unable to find image 'mongo:3.5' locally
B.5: Pulling from library/mongo
d13d02fa248d: Pull complete
bc8e2652ce92: Pull complete
3cc856886986: Pull complete
c319e9ec4517: Pull complete
b4cbf888f94: Pull complete
e42648c7b452: Pull complete
bfce988bf680: Pull complete
bd3ae40a4e35: Pull complete
dbdbb4a86824: Pull complete
d4624d0639de: Pull complete
Digest: sha256:82e1da5703ed53b03b8f7cac9e531b114d7493476e04b5d624e968f38bb97f07
Status: Downloaded newer image for mongo:3.5
ce711c6ddc679c292e7a793c677dc8aaa3b7b837ca0f5b19e4f689f6cd2e8754
PS C:\>
```

- MongoDB está preparado para escalar fácilmente de manera horizontal

```
PS C:\> docker run --name mongobongo -v /home/dac/Downloads/mongodb:/data/db -d mongo:3.5
Unable to find image 'mongo:3.5' locally
B.5: Pulling from library/mongo
d13d02fa248d: Pull complete
bc8e2652ce92: Pull complete
3cc856886986: Pull complete
c319e9ec4517: Pull complete
b4cbf888f94: Pull complete
e42648c7b452: Pull complete
bfce988bf680: Pull complete
bd3ae40a4e35: Pull complete
dbdbb4a86824: Pull complete
d4624d0639de: Pull complete
Digest: sha256:82e1da5703ed53b03b8f7cac9e531b114d7493476e04b5d624e968f38bb97f07
Status: Downloaded newer image for mongo:3.5
ce711c6ddc679c292e7a793c677dc8aaa3b7b837ca0f5b19e4f689f6cd2e8754
PS C:\>
```

- MongoDB está preparado para escalar fácilmente de manera horizontal

```
PS C:\> docker run --name mongobongo -v /home/dac/Downloads/mongodb:/data/db -d mongo:3.5
Unable to find image 'mongo:3.5' locally
B.5: Pulling from library/mongo
d13d02fa248d: Pull complete
bc8e2652ce92: Pull complete
3cc856886986: Pull complete
c319e9ec4517: Pull complete
b4cbf888f94: Pull complete
e42648c7b452: Pull complete
bfce988bf680: Pull complete
bd3ae40a4e35: Pull complete
dbdbb4a86824: Pull complete
d4624d0639de: Pull complete
Digest: sha256:82e1da5703ed53b03b8f7cac9e531b114d7493476e04b5d624e968f38bb97f07
Status: Downloaded newer image for mongo:3.5
ce711c6ddc679c292e7a793c677dc8aaa3b7b837ca0f5b19e4f689f6cd2e8754
PS C:\>
```

C. Inserción y consulta de datos

D. Comparación

E. Documental

La popularidad del término "base de datos orientada a documentos" o "almacén de documentos" ha crecido a

la par con el uso del término NoSQL.

Una base de datos NoSQL documental sirve para gestionar información orientada a documentos o datos semi-estructurados, que se almacenarán en forma de documento.

Normalmente los documentos se estructuran dentro de una serie de contenedores llamadas colecciones (en inglés “collections”) que son proporcionados por el sistema gestor documental. Estas utilizan punteros para el acceso de la información almacenada en la base de datos, esto se realiza sin la ayuda de una estructura entidad relación predeterminada, como en el caso de las bases de datos relacionales.

La búsqueda de la información se realiza con base al contenido del documento.

Las codificaciones más habituales de estos documentos suelen ser XML, YAML o JSON, pero también se pueden almacenar en formato Word o PDF. Las bases de datos de tipo documental son una forma moderna de almacenar datos en formato JSON en lugar de filas y columnas como en las bases de datos relacionales; y esto permite expresar los datos en su forma natural.

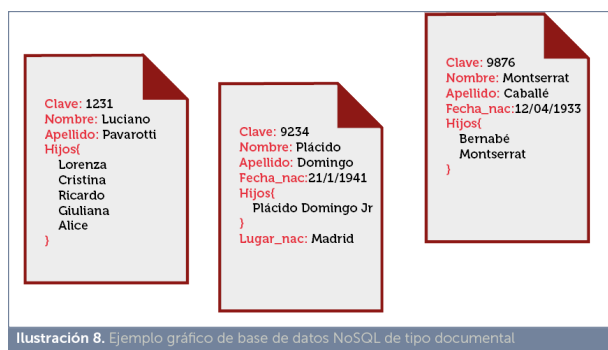


Ilustración 8. Ejemplo gráfico de base de datos NoSQL de tipo documental

Fundamentos del formato JSON :

- Pares de valores clave o atributos : Con el formato JSON se almacena en un par de valores clave. A estos pares se les llama a veces atributos. Las claves son cadenas simples y los valores pueden ser de cualquier tipo.
- Incrustación de objetos: Los valores incluidos en el par de valores clave también pueden ser otros objetos JSON, lo que permite crear una jerarquía de objetos. Colocar objetos JSON dentro de otro objeto JSON se denomina modelo de datos incrustados en base de datos documentales.
- Matrices: El formato JSON trabaja con un lenguaje de programación natural en todos los lenguajes de programación y estructuras de datos que son las matrices, por lo cual se admite almacenamiento de matrices como valores contra una clave.

```

1 {
2   {
3     "year": 2013,
4     "title": "Turn It Down, Or Else!",
5     "info": {
6       "directors": [ "Alice Smith", "Bob Jones" ],
7       "release_date": "2013-01-18T00:00:00Z",
8       "rating": 6.2,
9       "genres": [ "Comedy", "Drama" ],
10      "image_url": "http://ia-media-imdb.com/images/I/09ERUAT75797A37LUBR009AMUP988BL103F98EW7L3KQ7@_VI_SX480_.jpg",
11      "plot": "A rock band plays their music at high volumes, annoying the neighbors.",
12      "actors": [ "David Mattheuman", "Jonathan G. Heff" ]
13    }
14  },
15  {
16    "year": 2015,
17    "title": "The Big New Movie",
18    "info": {
19      "plot": "Nothing happens at all.",
20      "rating": 0
21    }
22  }
23 }
  
```

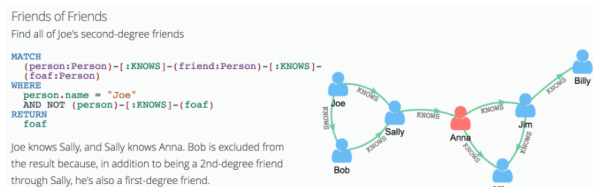
Factores que influyen en la aceptación de una BD NoSQL documental :

- Adopción de un nuevo paradigma: Este factor de adaptarse al concepto de NoSQL, resulta ser relativamente difícil, ya que la gran mayoría de compañías se encuentran más familiarizadas con las bases de datos relacionales.
- Curva de aprendizaje: al ser una tecnología relativamente nueva, esto representa para algunas personas un obstáculo en aprender la utilización de las bases de datos NoSQL
- Costo-beneficio: es uno de los factores más importantes, debido al hecho de comparar si los beneficios que se obtendrán son mayores a los costos de implementarlo.

F. Clave-Valor

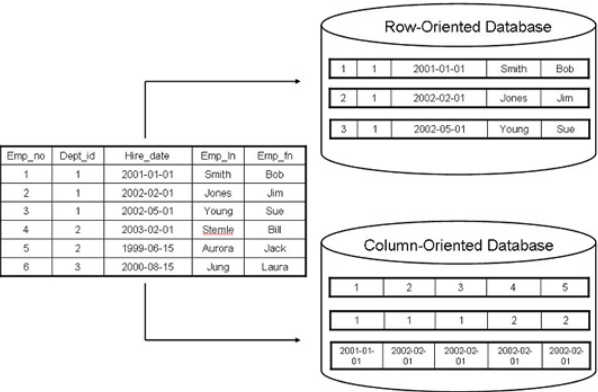
G. Grafos

- Estas bases de datos utilizan el modelo de Grafos
- Se especializan en relaciones
- Las podemos utilizar por ejemplo para guardar puntos de un camino, relaciones de amigos, familia, o cualquier tipo de dato que represente alguna relación
- Entre los motores más conocidos de este tipo se encuentra Neo4j



H. Tabular (Column-Store)

- Este tipo de familia de bases de datos está orientada a grandes cantidades de datos
- Lo datos son almacenados en columnas
- En una columna tiene múltiples datos
- Entre los motores más conocidos de este tipo se encuentra Cassandra o HBase



V. DISCUSIÓN Y CONCLUSIONES

- NoSQL permite el manejo de grandes volúmenes de datos y la posibilidad de tener un sistema distribuido.

I. Comparación entre BD Documental y Clave-Valor

	Documentales	Clave-Valor
Formatos	JSON y XML, son los más usados.	String, Numéricos, JSON y estructuras más complejas.
Identificadores	Cada documento consta de un ID único para facilitar métodos de indexación.	Los elementos únicos se identificarán mediante la tupla (bucket, clave).
Ventajas	<ul style="list-style-type: none">- Simplicidad- Eficiencia- Flexibilidad	Su organización
Aplicaciones	<ul style="list-style-type: none">- Datos de sensores provenientes de diferentes fabricantes.- Fichas de clientes con diferentes características.- Catálogos de inventario de productos para una tienda o fábrica.	<ul style="list-style-type: none">- Memorias caches de páginas web, donde la URL es la clave y el contenido el valor.- Logs de operaciones, con los timestamp como la clave y el contenido como valor.
Más usadas	<ul style="list-style-type: none">- MongoDB- CouchDB- CouchBase- MarkLogic	<ul style="list-style-type: none">- Riak KV- Redis

- Las características de las bases de datos NoSQL responden a las necesidades actuales de las diferentes organizaciones, por lo que son una alternativa debido a su capacidad y a la velocidad.